PAI : summary

2. BN

RV: $X : \Omega \to D$ $\quad P(X=x) = P(\{\omega \in \Omega : X(\omega) = x\})$

Bernoulli $D=\{H,T\}$ coin

Binomial $P(S=s) = \binom{n}{s} p^s (1-p)^{n-s}$

Categorical $D=\{1..m\}$ dice → Multinomial = #occurrences

Gaussian $\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ ⇒ Multivariate $\Sigma = \begin{pmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{21} & \sigma_2^2\end{pmatrix}$, $\mu = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}$

$\frac{1}{2\pi|\Sigma|} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$

$A \Rightarrow B = \neg A \vee B$

Conditional distrib $P(A|B) = \frac{P(A,B)}{P(B)}$ (new evidence can change belief)

Sum Rule $P(X_{1:i-1}, X_{i+1:n}) = \sum_{x_i} P(X_{1:i-1}, X_i = x_i, X_{i+1:n})$ → marginal

Product/chain Rule $P(X_{1:n}) = P(X_1)P(X_2|X_1)\dots P(X_n|X_{1:n-1})$

Prior $P(X)$ Bayes → Posterior $P(X|Y) = P(X)P(Y|X)$

Likelihood $P(Y|X)$ $\qquad \sum_x P(X=x)P(Y|X=x) \; \{ P(Y)$

High-dim: $P(X_{1:n}) \to 2^n - 1$ params

Indep RV $P(X_{1:n}) = \prod_{i=1}^n P(X_i) \to n$ params

marginals contrib $P(X_i) \to 2^{n-1}$ params; $P(X_i|X_j) \to 2^{n-2}$ params

*Conditional indep $P(T,V,C) = P(C|T,V) \cdot P(T|V) \cdot P(V)$ → $2^2 + 2 + 1$ params $\ll$

$P(C|T,V) \; \{2$ $\quad P(T|V) \; \{2$

$P(X,Y|Z) = P(X|Z)P(Y|Z)$ $\}$ $P \models X \perp Y | Z$ (Y doesn't say more than Z)
$P(X|Y,Z) = P(X|Z)$

Properties
- Symm $X \perp Y | Z \Rightarrow Y \perp X | Z$
- Decomp $X \perp Y,W | Z \Rightarrow X \perp Y | Z$
- Intersection $(X \perp Y | W,Z) \wedge (X \perp W | Y,Z) \Rightarrow X \perp Y,W | Z$ ($p > 0$)
- Contract $(X \perp Y | Z) \wedge (X \perp W | Y,Z) \Rightarrow X \perp Y,W | Z$
- Weak union $X \perp Y,W | Z \Rightarrow X \perp Y | Z,W$

Naïve Bayes model cause Y, Effects $X_{1:n}$ ⇒ $\forall i$ $X_i \perp X_{1..n/i} | Y$

$P(X_{1:n}, Y) = P(Y) \prod_{i=1}^n P(X_i|Y) \to 2n+1$ params

Inference $P(Y|X_i) = \dots = \frac{1}{Z} P(Y) \cdot P(X_i|Y)$.

3. BN + d separation

use conditional rather than joint parametrization → $\forall$ RV → specify $P(X_i|X_{A_i})$

$P(X_{1:n}) = \prod_i P(X_i|X_{A_i}) \to \sum_i 2^{|A_i|}$ params → $O(n)$ if $|A_i| \le k$

Bayesian Network → causal parametrization

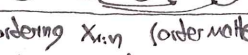
→ : immediate cause to immediate effects

$P(E,B,A,J,M) = P(E) \cdot P(B) \cdot P(A|E,B) \cdot P(J|A) \cdot P(M|A)$

DAG G → vertex = RV (BN structure) ← parents

(G,P) → P is CPT (cond. pr. distr) $P(X_i | Pa_{X_i})$ ⇒ $P(X_{1:n}) = \prod_i P(X_i|Pa_{X_i})$

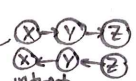
All prob can be described as BN (prod. rule)



World → BN
true distr P^* cond. ind $I(P^*)$ → (G,P) ⇒ $I(P)$

MAKE:
- pick ordering $X_{1:n}$ (order matters)
- $\forall i$ → find min subset A of $\{1..i-1\}$ st $X_i \perp X_{\bar{A}} | X_A$ $\quad \bar{A} = \{1..i-1\} \setminus A$
→ learn $P(X_i|X_A)$

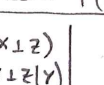
3 nodes



indirect causal effect / indirect evidential effect

common cause (∝)

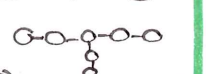
$\neg(X \perp Z)$ $\quad (X \perp Z | Y)$



$X \perp Z$ $\quad \neg(X \perp Z | Y)$

V-structure common effect (β)

Identification of CI: Active trails → see example



Active trails: undirected path is at for observed RV O ⇔
$\forall$ triplets on path $X-Y-Z$
(∝) → Y unobserved ($\notin O$)
(β) → Y or its descendant observed ($\in O$)

dsep
$\forall X_i, X_j$ → $\nexists$ a.t. for O → they are d-separated by O $\quad$ d-sep$(X_i, X_j | O)$

A, B are dsep given O ⇔ $\forall X \in A, Y \in B$ dsep$(X, Y | O)$

Independence dsep $(X;Y|Z) \to X \perp Y | Z$ (doesn't exist a.t.)

(converse doesn't hold (but for almost all))

4. BN - exact inference

dsep Alg: check every path → exponential :(to check dsep$(X;Y|Z)$
1. find all reachable from X 2. mark Z & ancestors 3. BFS from X, stop if blocked

Queries:
- conditional distrib
- maximization (MPE, most prob expl) $\quad \operatorname{argmax}_{e,b,a} P(E,B,A|J,M)$ (all vars)
(MAP, max of post) $\quad \operatorname{argmax}_{e,b} \sum_a$ " (some var)

#P compl

Hardness NP complete →

Inference: exploit structure to perform exact inference

Variable Elim: see example
↳ push sums through
↳ sum out vars

$p(x_4) \sum_{x_2} p(x_2|x_4) \sum_{x_3} p(x_3|x_2) \sum_{x_5} p(x_4|x_3)\, p(x_5|x_4)$
$g_2(x_4, x_5)$ $\quad g_3(x_2, x_5)$ $\quad g_4(x_3, x_5)$

V.E. Alg ▷ BN, query $P(X|E=e)$ evidence
1. choose ordering $X_{1:n}$
2. setup factors $f_i := P(X_i|Pa_i)$
3. for $i = 1:n$ $\quad X_i \notin \{X, E\}$
- collect/multiply f_j that include X_i
- generate by marginalization
4. renormalize $P(x,e) \to P(x|e)$

$g = \sum_{x_i} \prod_j f_j$ (add it to factors)

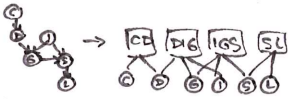
(order matters)

Polytree: Dag is polytree ⇔ removing directions yields tree
↳ V.E.: pick root → orient edges → remove nodes in top. order
permutation π → $x_i \in desc(x_j)$ ⇒ $\pi(i) > \pi(j)$

5. BN & belief propagation

Answering multiple queries: ⇔ find $P(X_i|X_n=T)\ \forall i$ naïv: V.E $\forall i$ → $O(n^2)$
⇒ reuse computation!
message passing on factor graph

Factor graph (bip graph, vars & factors)
(all CPT have to be assigned)



Remark: suppose $P(X_{1:n}) = \frac{1}{Z} \prod_{i=1}^m \psi(X_{A_i})$ $\quad A_i \in \{1..n\}$ $\quad \psi$ = factors

$Z = \sum_x \prod_i \psi_i(x_{A_i})$ normalization

Image Segmentation
X_i: label (bg/fg)
Y_i: greyscale

$P(x) = \frac{1}{Z} \prod_i \phi(x_i) \prod_{(j,k) \in E} \psi(x_j, x_k)$

$\phi(x_i) = \exp\left(-\frac{(y_i - \mu_{x_i})^2}{2\sigma_{x_i}^2}\right)$

$\psi(x_i, x_j) = \exp\left(-\beta(x_i - x_j)^2\right)$

want to be close to mean $\quad$ penalize discontinuity

Incorporate observation: multiply with indicator fct $\quad [C=T](c) = \begin{cases} 1 & c=T \\ 0 & \text{else} \end{cases}$
- renormalize marginals

$\sum_{abc} P(a,b,c|D=T) = \sum_{abcd} P(a,b,c,d)\,[D=T](d)$

Sum-Product msg passing

node v, factor u $\quad$ init: all = 1

$\mu_{v \to u}(x_v) = \prod_{u' \in N(v) \setminus \{u\}} \mu_{u' \to v}(x_v)$

$\mu_{u \to v}(x_v) = \sum_{x_{u'} = x_v} f_u(x_u) \prod_{v' \in N(u) \setminus v} \mu_{v' \to u}(x_{v'})$

Belief prop. Alg
- init all as uniform
- until converged → {- pick some ordering - update msg according to - break if changed by $\le \varepsilon$

Hope: @end $P(X_v = x_v) = \frac{1}{Z} \prod_{u \in N(v)} \mu_{u \to v}(x_v)$

$P(X_v = x_v)$ (marginals) $\propto f_v(x_v) \prod_{u \in N(v)} \mu_{u \to v}(x_v)$

Poly trees: BP is exact! (choose root, msg leaves → root → leaves)

Approximate inference with loops → no convergence or incorrect

Loopy BP (multiplies some factors multiple times) → over confident
oscillation more severe ⇔ potentials more 'deterministic'

6. Approx inference via Sampling

Variable Elim. MPE same as V.E: ∗ - generate by maximization
$g_i = \max_{x_i} \prod_j f_j$
4. $i = n..1$ $\quad \hat{x}_i = \operatorname{argmax}_{x_i} g_i(x_i, \hat{x}_{i+1:n})$

Msg passing replace (Σ) by max

Retrieving MAP $P_{max}(X_v = x_v) = \max_{x \sim x_v} P(x)$

Sampling Based inference (stochastic approx) → converges but slow

Marginal as expectation: marginal as expectation of indicator
$P(x_i) = \sum_{x_{j \neq i}} P(x_{1..n}) = \sum_{x_{1..n}} P(x_{1..n})[x_i = x_i] = E_P[x_i = x_i]$

Sample approx of expect $\quad x_1, \dots, x_N \sim P(x)$ $\quad$ strong law large numb
$E_P[f(x)] = \lim_{N \to \infty} \frac{1}{N} \sum f(x_i)$

Hoeffding's ineq $P(|E_p[f(x)] - \frac{1}{N}\sum f(x_i)| > \varepsilon) \le 2\exp(-2N\varepsilon^2/c^2) \le \delta$ ($f \in [0,c]$)

error $< \varepsilon$ with prob $\ge 1-\delta$ → $N \ge \frac{c^2}{2\varepsilon^2}\ln 2/\delta$

Sampling Bernoulli $x = \begin{cases} T \\ F \end{cases} \Leftrightarrow u \le p$, $u \sim Unif[0,1]$

Multinomial $X = g(u)$, $g: [0,1] \to \{1..k\}$

Forward sampling (conditioned on previous result)

Monte Carlo top.sort $i = 1..n$ $x_i \sim P(X_i \mid x_{1..i-1})$

Marginals $P(X_i = x_i) \approx \frac{count(X_i = x_i)}{N} = \hat{P}(X_i)$

Conditionals $P(X_i \mid X_j) = count(X_i, X_j)/count(X_j)$

Rejection sampling → problem if rare event ⇒ sample from posterior distrib!

Relative error $P(\hat{P}(x) \in P(x)(1 \pm \varepsilon)) \le 2\exp(-NP(x)\varepsilon^2/3)$

7. MC MC

Posterior distrib sampling $P(x_A \mid x_B^{ob}) = \frac{1}{Z} P(x_A, x_B^{ob})$ ($Q(x_A)$)

Normalize intractable: sample from $P(x)$?

MC: create Markov chain w/ stationary distribution $P(x)$

MC: $x_1 \to x_2 \to \dots \to x_m$ RV $X_{1..N}$

Prior $P(X_A)$, Transitions $P(X_{t+1} \mid X_t)$

Ergodic $\Leftrightarrow \exists t$ st every state reachable from anywhere in t steps

Stationary distrib $\pi(X) > 0$ st $\forall x \lim_{N\to\infty} P(X_N = x) = \pi(x)$ (indep of X_1)

sample long enough → sample similar to π

MC MC ▷ $Q(x)$ unnormalized → design MC with $\pi(x) = \frac{1}{Z}Q(x)$

Detailed Balance equation $\frac{1}{Z}Q(x)P(x' \mid x) \overset{DB}{=} \frac{1}{Z}Q(x')P(x \mid x')$ | MC satisfies if $\forall x, x'$

Designing it 1) proposal $R(x' \mid x)$ 2) acceptance

- suppose $X_t = x$
- $\alpha = \min(1, Q(x')R(x \mid x')/Q(x)R(x' \mid x))$
- $x_{t+1} = \begin{cases} x' & \text{prob } \alpha \text{ (proposed state)} \\ x & \text{prob } 1-\alpha \text{ (old state)} \end{cases}$

Gibbs sampling (practical)

- initial assignment $\forall x$ $x^{(0)}$
- fix observed $X_B = x_B$
- for $t = 0..$
  - set $x^{(t)} = x^{(t-1)}$
  - foreach $X_i \notin B$ → v_i = val of all $x^{(t)}$ except x_i; sample $x_i^{(t)}$ from $P(X_i \mid v_i)$

Insight: sampling efficient! $P(X_i \mid v_i) = \frac{1}{Z}\prod_{j:i \in f_j} f_j(v_i)$

joint sample depends on previous two → Low Iorge Numb does not apply

Ergodic theorem ▷ $X_{1..N}$ ergodic MC, state space D, f fct on D ⇒ $\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} f(x_i) = \sum_{x \in D} \pi(x) f(x)$

let MC 'burn in' → ignore first to samples

$E[f(x)] \approx \frac{1}{T - t_0}\sum_{t=t_0+1}^{T} f(x^{(t)})$

static { Sum: exact marg • VE, BP; loopy net • loopy BP • variational inf • Gibbs

8. SEQUENTIAL MODELS

Dynamic models $X_{t-2} \to X_{t-1} \to X_t \to \dots$

markov assump. $X_{1:t-1} \perp X_{t+1:T} \mid X_t$

stationary assumpt. $P(X_{t+1} = x \mid X_t = x')$ not dep on t

HMM • speech recogn,..

Kalman • cars/missiles..

HMM: Categorical / Kalman: gaussian

hidden state → observ

- Filtering (state estimation) $P(X_t \mid u_{1:t})$
- Prediction $P(X_{t+k} \mid u_{1:t})$
- Smoothing $P(X_t \mid u_{1:T})$ $1 \le t \le T$
- most prob expl $\arg\max_{x_{1:T}} P(x_{1:T} \mid u_{1:T})$

Efficient exp - independent of t

Bayes Filt: computed $P(X_t \mid u_{1:t})$ → compute $P(X_{t+1} \mid u_{1:t+1})$

Conditioning: $P(X_t \mid u_{1:t}) = \frac{1}{Z}P(X_t \mid u_{1:t-1}) \cdot P(u_t \mid x_t, u_{1:t-1})$

Predict: $P(X_{t+1} \mid u_{1:t}) = \sum_{x_t} P(X_{t+1} \mid x_t, u_{1:t}) P(x_t \mid u_{1:t})$

Prior $P(X_1)$ wise

Smoothing / MPE: HMM is polytree → use sum-product / fwd-bwd (= max-prod / Viterbi) for smoothing / MPE

Kalman (continuous domain) — everything is gaussian

$X_{t+1} = X_t + \Delta X_t'$ (true interval)

$P(X_{t+1} = x_{t+1} \mid X_t = x_t, X_t' = x_t') = \mathcal{N}(x_t + \Delta x_t', \sigma^2)$

Multivariate gaussian thm. $X_W = [X_A, X_B] \sim \mathcal{N}(\mu_W, \Sigma_{WW})$ ↳ consider subset $A = \{i_1..k\} \to X_A \sim \mathcal{N}(\mu_A, \Sigma_{AA})$

→ $P(X_A \mid X_B = x_B) = \mathcal{N}(\mu_{A|B}, \Sigma_{A|B})$ | $\mu_{A|B} = \mu_A + \Sigma_{AB}\Sigma_{BB}^{-1}(x_B - \mu_B)$, $\Sigma_{A|B} = \Sigma_{AA} - \Sigma_{AB}\Sigma_{BB}^{-1}\Sigma_{BA}$

Multiple of Gaussian is Gaussian $X \sim \mathcal{N} \to \Pi X = Y \sim \mathcal{N}(\Pi\mu_W, \Pi\Sigma_{WW}\Pi^T)$

Sum is also $X + X' = Y \sim \mathcal{N}(\mu_W + \mu_W', \Sigma_{WW} + \Sigma_{WW}')$

Kalman = Gaussian HMM

assume linear model

$X_{t+1} = FX_t + \varepsilon_t \sim \mathcal{N}(0, \Sigma_X)$ ← motion model $P(X_{t+1} \mid X_t)$

$Y_t = HX_t + \eta_t \sim \mathcal{N}(0, \Sigma_Y)$ ← sensor model $P(Y_t \mid X_t)$

Bayesian filtering KF $P(x_t \mid y_{1:t}) = \frac{1}{Z}P(x_t \mid y_{1:t-1})P(y_t \mid x_t)$ correction

$P(x_{t+1} \mid y_{1:t}) = \int P(x_{t+1} \mid x_t)P(x_t \mid y_{1:t})\,dx_t$ prediction

Ex random 1D walk $x_{t+1} = x_t + \varepsilon_t \sim \mathcal{N}$

$\mu_{t+1} = \frac{\sigma_q^2\mu_t + (\sigma_t^2 + \sigma_x^2)y_{t+1}}{\sigma_t^2 + \sigma_x^2 + \sigma_q^2}$ | $x_{t+1} = x_{t+1} + \mu_{t+1} \sim \mathcal{N}$

9. MDPs

Kalman update Trans. model $P(x_{t+1} \mid x_t) = \mathcal{N}(x_{t+1} \mid Fx_t, \Sigma_x)$ Sensor model $P(y_t \mid x_t) = \mathcal{N}(y_t \mid Hx_t, \Sigma_y)$ | KL

$\mu_{t+1} = F\mu_t + K_{t+1}(y_{t+1} - HF\mu_t)$ (observ. expect)

$\Sigma_{t+1} = (I - K_{t+1}H)(F\Sigma_t F^T + \Sigma_x)$

Kalman gain $K_{t+1} = \dots$

Dynamic BN (BN → EBN) .. replicate over time

Flow of influence lots of dependencies over time

slice, temporal edges

Particle filtering 1. propagate 2. weight (based on observ) 3. resample

represent distrib with particles $P(x) \approx \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}(x)$ | $E_{x\sim\delta_{x'}}[f(x)] \approx f(x')$

$\Delta P(x_t \mid y_{1:t}) = \frac{1}{N}\sum \delta_{x_i,t}$ 1. $x_i' \sim P(X_{t+1} \mid x_{i,t})$ 2. $w_i = \frac{1}{Z}P(y_{t+1} \mid x_i')$ 3. $x_{i,t+1} \sim \sum_{i=1}^{N} w_i \delta_{x_i'}$ $\mathcal{N}(x_i', 0)$

PF in DBN each particle represents a realization of slice S_t, i.e. $x_{t,i} = [a_{t,i}, b_{t,i}, \dots]$

SLAM = simultaneous localization and mapping

Probabilistic planning → "how to control robot to maximize reward"

MDP state X, action A, transition prob $P(x' \mid x, a)$, reward $r(x,a)$ | planning: policy $\pi: X \to A$ → $P(X_{t+1} = x' \mid X_t = x, \pi) = P(x' \mid x, \pi(x))$

expected value $J(\pi) = E[r(x_0, \pi(x_0)) + \gamma r(x_1, \pi(x_1)) + \gamma^2 \dots]$ γ discount factor $\in [0,1]$ (prefer immediate reward)

Value function $V^\pi(x) = J(\pi \mid X_0 = x) = \dots r(x, \pi(x)) + \gamma \sum_{x'} P(X_1 = x' \mid X_0 = x) \cdot E[\sum_{t=0}^{\infty} \gamma^t r(X_{t+1}, \pi(X_{t+1})) \mid X_1 = x'] = \dots V^\pi(x')$

$V^\pi = \begin{bmatrix} V^\pi(1) \\ \vdots \\ V^\pi(n) \end{bmatrix}$ $T^\pi = \begin{pmatrix} P(1 \mid 1, \pi(1)) & \dots \\ & \ddots \\ & P(n \mid n, \pi(n)) \end{pmatrix}$ ⇒ $V^\pi = r^\pi + \gamma T^\pi V^\pi$ ⇒ $V^\pi = (I - \gamma T^\pi)^{-1} r^\pi$

$\pi^* = \arg\max_\pi J(\pi)$ $O(n^m)$ → Greedy policy wrt V: $\pi_V(x) = \arg\max_a r(x,a) + \gamma \sum_{x'} P(x' \mid x, a) V(x')$

Bellman: Policy opt ⇔ greedy wrt induced V^π $V^*(x) = \max_a r(x,a) + \gamma \sum_{x'} P(x' \mid x, a) V^*(x')$

converges bc. contraction

10. Learning BNs

Policy iteration start π random, until converged: $V^\pi(x) \to \pi_G$ wrt V^π | set $\pi^* = \pi_G$ $O(n^2 m)$

$\forall x$ $V^{\pi_{t+1}} \ge V^{\pi_t}$

Compute V^* with DP: $V_t(x) = \max$ reward if world ends in t more steps $V_0(x) = \max_a r(x,a)$, $V_t(x) = \max_a r(x,a) + \gamma \sum_{x'} P(x' \mid x, a) V_0(x')$

$Q_t(x,a) = r(x,a) + \gamma \dots$ → $\exists$ fixpoint & converges

Contraction: $B: \mathbb{R}^n \to \mathbb{R}^n$ Thm: $\|BV - BV'\|_\infty \le \gamma \|V - V'\|$

Tradeoff: value vs policy iter → exact/polynomial time

MDP = controlled MC $P(X_{t+1} \mid X_t, A)$

POMDP = " HMM

POMDP: states: $B = \Delta(\{1..n\}) = \{b : \{1..n\} \to [0,1], \sum b_i = 1\}$

Actions: same

Trans: $P(b_t' \mid b_t) = \sum_{x=1}^{n} P(b_t' \mid X_t = x) b_t(x)$

State update = belief state MDP → new state is belief | $P(Y_t \mid b_t)$

Approx sol: reduce & sample

Bayes filtering

Policy gradient methods

parametric form: $\pi(b) = \pi(b; \theta)$ param → compute $E[J]$ by sampling → $\theta^* = \arg\max J(\theta)$

11. Learning structure

Learning BN from data

Parameter learning: Bernoulli RV $\theta^* \in \arg\max P(D \mid \theta) = \arg\max \sum \log P(f(i) \mid \theta)$ … $\Rightarrow \theta^* = \frac{n_+}{n_+ + n_-}$

Estimate CDP $\frac{count(X \mid Y)}{count(Y)}$ { given $D = \{x_{1..N}\}$, MLE $P(x_{1..N}) = \prod P(x_i \mid Pa_i, \theta_{i \mid Pa_i})$

MLE for BN

Alg for BN MLE: $\hat{\theta}_{x_i \mid Pa_i} = \frac{count(x_i, Pa_i)}{count(Pa_i)}$

Pseudo-counts: prior assumption about param (eg. fair coin) $\theta_{T=c} = \frac{count(T=c) + \alpha_c}{N + \alpha_c + \alpha_{\bar{c}}}$

(beta prior) → $Beta(\theta, \alpha_c, \alpha_{\bar{c}}) = \frac{1}{B(\alpha_c, \alpha_{\bar{c}})}\theta^{\alpha_c - 1}(1-\theta)^{\alpha_{\bar{c}} - 1}$

suppose observation: $P(\theta, D) \propto P(\theta \mid \alpha_c, \alpha_{\bar{c}})P(n_c, n_{\bar{c}} \mid \theta) = \frac{1}{Z(\alpha_c, \alpha_{\bar{c}}, n_c, n_{\bar{c}})}\theta^{\alpha_c + n_c - 1}(1-\theta)^{n_{\bar{c}} + \alpha_{\bar{c}} - 1}$

Think of learning as inference

in augmented BN with variables replicated per observation

infer $P(\theta, \theta_1, \dots \mid data)$ eg $\hat{\theta}_{1..} = \arg\max P(\theta..)$ MPE

Score based structure learning score $S(G, D)$ structure, data

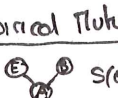
$G^* = \arg\max_G S(G;D)$ $S(G;D) = \max_\theta \log P(D \mid \theta, G) = \max_\theta N\sum_{i=1}^{n} \hat{I}(X_i; Pa_{X_i}) + const$

MLE score

Mutual info $I(X_i, X_j) = \sum_{x_i, x_j} P(x_i, x_j) \log \frac{P(x_i, x_j)}{P(x_i)P(x_j)}$ $= 0 \Leftrightarrow X_i, X_j$ indep

$I(A,B) = I(B,A)$, $B \subset C \Rightarrow I(A,B) \le I(A,C)$

Empirical Mutual info $\hat{P}(x_i, x_j) = \frac{\text{Count}(x_i, x_j)}{N} \to \hat{I}$ uses $\hat{P}$

ex  $S(G,D) = \sum \hat{I}(X, P_{Xi}) = \hat{I}(E,\emptyset) + \hat{I}(B,\emptyset) + \hat{I}(A;E|B) + \dots$

optimal solution: complete graph :(

Sol: prior/constraint on params/structure ex $S_{BIC}(G) = \sum_{i=1}^{N} I(x_i, P_{x_i}) - \frac{\log N}{2N}|G|$ Transfer learning/handle context

BIC is consistent (correct as $|N| \to \infty$ = # samples)

Local search:... → MST

Finding opt tree ▸ $G=(V,E)$, $w(e) = I(X_i, X_j)$ find tree with min $\sum w_e$

Chow-Liu ALG $\forall X_i, X_j$ comp $\hat{P}(X_i, X_j) = \frac{\text{count}(X_i, X_j)}{N}$ → comp mutual info $\hat{I}$
build weighted graph, find MST, pick root, orient edges away → have BN

12. Intro to RL (reinf. learning)

Learning from actions, actions change state of world

Different from supervised: data is not iid

Exploration/Exploitation dilemma: gather data or use what know already
• random action? • best action wrt knowledge?

1) Model Based RL (approx DP): learn MDP : estimate $P(x'|x,a)$, $r(x,a)$
optimize policy

Experiences: $x_1, a_1, r_1, x_2, \dots$ $D = \{(x_1, a_1, r_1, x_{1+1})\}_i^t$; $\approx \frac{\text{Count}(X_{t+1}=x', X_t=x, A=a)}{\text{Count}(X_t=x, A=a)} \overset{N\to\infty}{=} \frac{1}{N_{x,a}} \sum R_T$

Greedy: Prob ε_t, pick random, else pick best
• if ε_t satisfies RM, then converge to optimal policy
$(\varepsilon_t \to 0, \sum_t \varepsilon_t \to \infty)$ eg $\varepsilon_t = \frac{1}{t}$

Optimism in face of uncertainty $r(x,a)$? ← R_{max}
(no need to select explor/exploit) $P(x'|x,a)$? ← use $x' = x^*$ (very bold)

Thm: every T timesteps: R_{max} either: near optimal / visit unknown state

R_{max} alg: in: x_0, γ | init: x^* added to MDP, set $r(x,a) = R_{max}$, $P(x^*|x,a) = 1$ $\forall$
repeat: choose optimal policy for r, P
Execute π, update r, P, recompute π

How many samples? → Hoeffding-Chernoff
Thm: with prob $1-\delta$, R_{max} reaches ε-optimal policy in $\mathcal{O}(\text{pol}(|X|, |A|, T, \frac{1}{\varepsilon}, \log\frac{1}{\delta}))$
Model-based RL: memory $\mathcal{O}(X^2 A)$, time $\mathcal{O}(\text{poly}|X||A|)$

2) Model-Free RL: optimal $V^*(x) \to$ optimal π^*

$V^*(x) = \max_a Q^*(x,a)$, $Q^*(x,a) = r(x,a) + \gamma \sum_{x'} P(x'|x,a) V^*(x')$ ← estimate from data

Q-learning: initial estimate, then $Q(x,a) \leftarrow (1-\alpha_t) Q(x,a) + \alpha_t (r + \gamma \max_{a'} Q(x', a'))$
approx by: old estimated V^*, sample $x' \sim P(x'|x,a)$ (old / new)

if α_t satisfies RM ($\sum \alpha_t = \infty$, $\sum \alpha_t^2 < \infty$) converges to Q^* (with random actions)

Smilar to R_{max}: init $Q(x,a) = \frac{R_{max}}{1-\gamma} \prod_{t=1}^{T} (1-\alpha_t)^{-1}$ | Memory: $\mathcal{O}(|X||A|)$
Comput: $\mathcal{O}(A)$

Challenges: X,A exponential (chess) or continuous, risk of exploration

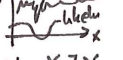
Parametric Q fct: $Q(x,a;\theta) = \theta^T \phi(x,a)$ = deepNN (optimize via GD)
$L(\theta) = \sum (r_t + \max_a \gamma Q(x', a'; \theta^{old}) - Q(x,a;\theta))^2$

13. Advanced RL

⊕ Simple, converges, train w/o policy
⊖ lot of exploration/need fully observable

Policy search methods: learn parametrized fct $\pi(x;\theta)$ → expected reward $J(\theta)$ by
rollouts: $\theta^* = \arg\max_\theta J(\theta)$

Learning to optimize bbox
build probabilistic model: acquisition function $D \ni x_t \to [f] \to y_t = f(x_t) + \epsilon_t$ | want: $f(x_t) \to f(x^*)$

Gaussian Processes $f(x)$  Likelihood: $P(\text{data}|f)$
Posterio: $P(f|\text{data})$

∞-D gaussian
a set of RV, indexed by X. $\forall x \in X \exists Y_x$. $\mu: X \to \mathbb{R}$, $K: X \times X \to \mathbb{R}$ st $\forall A \subseteq X: Y_A \sim \mathcal{N}(\mu_A, \Sigma_{AA})$
K: symmetric, positive

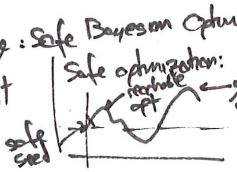
eg. Squared exp. kernel: $K(x,x') = \exp(-(x-x')^2/h^2)$ kernel
$= \exp(-|x-x'|/h)$
Exponential
$= x^T x'$
Linear:
$= \phi(x)^T \phi(x')$
Features

Making predictions with GP
Suppose $P(f) = GP(f; \mu, K)$, make observation, recompute μ, K with closed form
Focus exploration where upper confidence bound ≥ best lower bound

$x_t = \arg\max_{x \in D} \mu_{t-1}(x) + \beta_t \sigma_{t-1}(x)$

Thm: for bounded $\|f\|_k$, by selecting $\beta_t = \mathcal{O}(\gamma_t \log^3 t)$: $\min\{f(x^*) - f(x)\} = \mathcal{O}^*(\frac{\gamma_T}{\sqrt{T}})$
$\gamma_T = \max_{|A| \le T} I(f; y_A)$

Application: state automate controller tuning: Safe Bayesian Optimization
$\max_x f(x)$ st $g(x) \ge \tau$ ← safety constraint
reach performance: $\max_x f(x)$

Safe optimization: 
reachable opt
global optimum
safe seed

Safe opt: Thm: with suitable conditions on kernel,
$\exists$ fc $\tau(\varepsilon,\delta)$ st with $P \ge 1-\delta$: • safe opt never makes unsafe decision,
• after T steps reaches ε-opt reachable pt

Exploration: Virtual vs Physical: $\hat{f}(x) = f(x) + \mathcal{N}(0)$, $f(x)$
How many bits/cost